

# SMART AGRICULTURE CROP ANALYTICS



# Data Sourcing and Collection

We collected data on various parameters such as state names, crop types, soil nutrients, pH levels, rainfall, temperature, and crop yields from agricultural databases and direct surveys. This comprehensive data helps in analyzing and improving crop production efficiency. 🌾📊

[11]:	Unnamed: 0	State_Name	Crop_Type	Crop	N	P	K	pH	rainfall	temperature	Area_in_hectares	Production_in_tons	Yield_ton_per_hec
0	0	andhra pradesh	kharif	cotton	120	40	20	5.46	654.34	29.266667	7300.0	9400.0	1.287671
1	1	andhra pradesh	kharif	horsegram	20	60	20	6.18	654.34	29.266667	3300.0	1000.0	0.303030
2	2	andhra pradesh	kharif	jowar	80	40	40	5.42	654.34	29.266667	10100.0	10200.0	1.009901
3	3	andhra pradesh	kharif	maize	80	40	20	5.62	654.34	29.266667	2800.0	4900.0	1.750000
4	4	andhra pradesh	kharif	moong	20	40	20	5.68	654.34	29.266667	1300.0	500.0	0.384615
...	...	...	...	...	...	...	...	...	...	...	...	...	...
99844	99844	west bengal	rabi	wheat	60	30	30	6.70	152.54	22.280000	2013.0	5152.0	2.559364
99845	99845	west bengal	summer	maize	80	40	20	5.68	182.50	29.200000	258.0	391.0	1.515504
99846	99846	west bengal	summer	rice	80	40	40	5.64	182.50	29.200000	105.0	281.0	2.676190
99847	99847	west bengal	rabi	rice	80	40	40	5.42	152.54	22.280000	152676.0	261435.0	1.712352
99848	99848	west bengal	rabi	sesamum	30	15	30	6.54	152.54	22.280000	244.0	95.0	0.389344

# Data Cleaning

We deleted the "Unnamed: 0" column to streamline the dataset, ensuring only relevant information is included to improve data clarity and accuracy and to reduce complexity, making further analysis easier.

```
[13]: ds.drop(columns= ["Unnamed: 0"])
```



	State_Name	Crop_Type	Crop	N	P	K	pH	rainfall	temperature	Area_in_hectares	Production_in_tons	Yield_ton_per_hec
0	andhra pradesh	kharif	cotton	120	40	20	5.46	654.34	29.266667	7300.0	9400.0	1.287671
1	andhra pradesh	kharif	horsegram	20	60	20	6.18	654.34	29.266667	3300.0	1000.0	0.303030
2	andhra pradesh	kharif	jowar	80	40	40	5.42	654.34	29.266667	10100.0	10200.0	1.009901
3	andhra pradesh	kharif	maize	80	40	20	5.62	654.34	29.266667	2800.0	4900.0	1.750000
4	andhra pradesh	kharif	moong	20	40	20	5.68	654.34	29.266667	1300.0	500.0	0.384615
...	...	...	...	...	...	...	...	...	...	...	...	...
99844	west bengal	rabi	wheat	60	30	30	6.70	152.54	22.280000	2013.0	5152.0	2.559364
99845	west bengal	summer	maize	80	40	20	5.68	182.50	29.200000	258.0	391.0	1.515504
99846	west bengal	summer	rice	80	40	40	5.64	182.50	29.200000	105.0	281.0	2.676190
99847	west bengal	rabi	rice	80	40	40	5.42	152.54	22.280000	152676.0	261435.0	1.712352
99848	west bengal	rabi	sesamum	30	15	30	6.54	152.54	22.280000	244.0	95.0	0.389344

99849 rows × 12 columns

# Data Summary

- Nutrients (N, P, K): Average values show significant variations, with nitrogen (N) levels being the highest.
- pH Levels: The average pH is 5.64, with a range from acidic (3.82) to neutral (7.00).
- Environmental Factors: Rainfall averages at 701 mm, with a wide distribution up to 3322 mm. Temperature averages at 26.68°C.
- Crop Metrics:
  - Area in Hectares: Ranges significantly, showing the diversity in crop-growing scales.
  - Production in Tons and Yield per Hectare: Displaying variability, with yields reaching up to 9801 tons per hectare in some cases.

	count	mean	std	min	25%	50%	75%	max
N	99849.0	69.816823	39.571469	10.000000	50.000000	75.000000	80.000000	1.800000e+02
P	99849.0	41.593656	15.056508	10.000000	40.000000	40.000000	60.000000	1.250000e+02
K	99849.0	42.037827	28.430263	10.000000	20.000000	30.000000	50.000000	2.000000e+02
pH	99849.0	5.643624	0.505283	3.820000	5.360000	5.540000	5.960000	7.000000e+00
rainfall	99849.0	701.151085	604.701552	3.274569	157.310000	579.750000	1110.780000	3.322060e+03
temperature	99849.0	26.684154	4.851214	1.180000	23.106000	27.333333	29.266667	3.534667e+01
Area_in_hectares	99849.0	16476.585668	43604.268231	0.580000	130.000000	1010.000000	8099.000000	7.263000e+05
Production_in_tons	99849.0	37762.912080	122244.670727	0.000000	162.000000	1506.000000	14395.000000	3.530571e+06

# Data Cleaning and Preprocessing

## 1 Outlier Detection

Identified and removed outliers to maintain data integrity.

## 2 Data Normalization

Scaled numerical features to ensure consistency in the dataset.

## 3 Feature Engineering

Created new features based on existing data to improve model performance.



# Data Cleaning and Preprocessing

## 1 Handling Missing Values

Imputed missing values using mean for numerical data and mode for categorical data.

## 2 Removing Duplicate Values

Scaled numerical features to ensure consistency in the dataset.

## 3 Encoding

Created new features based on existing data to improve model performance.



# STATISTICAL ANALYSIS

1

## Correlation Matrix

Created a correlation matrix to examine relationships between variables, helping to identify patterns and significant relationships.e.

2

## Heatmaps

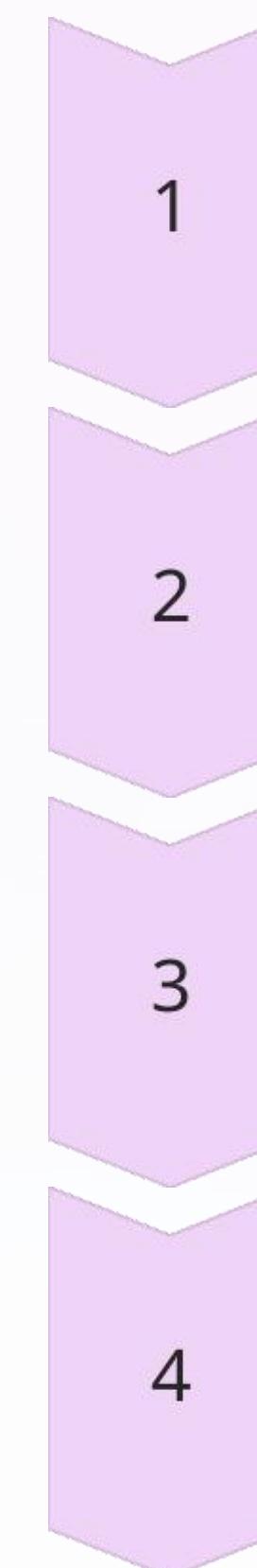
Used heatmaps to visually represent the correlation matrix, highlighting strong positive and negative correlations.

Correlation Matrix Heatmap

	State_Name	Crop	N	P	K	pH	rainfall	temperature	Area_in_hectares	Production_in_tons	Yield_ton_per_hec	Crop_Type_kharif	Crop_Type_rabi	Crop_Type_summer	Crop_Type_whole year
State_Name	1.00	0.00	-0.03	-0.03	-0.03	0.00	-0.24	0.13	0.06	0.10	0.01	0.03	0.01	0.04	-0.07
Crop	0.00	1.00	-0.22	-0.15	0.01	-0.08	-0.15	-0.04	0.18	0.14	-0.03	0.06	0.11	0.05	-0.20
N	-0.03	-0.22	1.00	0.31	0.48	-0.18	0.11	-0.01	0.04	0.10	0.09	-0.10	-0.08	-0.01	0.20
P	-0.03	-0.15	0.31	1.00	0.25	-0.25	0.11	-0.06	-0.06	-0.01	0.08	-0.07	-0.05	0.00	0.13
K	-0.03	0.01	0.48	0.25	1.00	-0.25	0.36	-0.09	-0.12	-0.01	0.09	-0.34	-0.17	-0.10	0.61
pH	0.00	-0.08	-0.18	-0.25	-0.25	1.00	-0.02	0.01	0.06	0.07	0.01	-0.01	0.02	-0.02	-0.00
rainfall	-0.24	-0.15	0.11	0.11	0.36	-0.02	1.00	0.05	-0.14	-0.10	0.03	0.16	-0.59	-0.26	0.58
temperature	0.13	-0.04	-0.01	-0.06	-0.09	0.01	0.05	1.00	-0.02	-0.03	0.00	0.43	-0.60	0.24	-0.01
Area_in_hectares	0.06	0.18	0.04	-0.06	-0.12	0.06	-0.14	-0.02	1.00	0.74	-0.02	0.08	0.14	-0.07	-0.19
Production_in_tons	0.10	0.14	0.10	-0.01	-0.01	0.07	-0.10	-0.03	0.74	1.00	0.03	-0.00	0.11	-0.05	-0.09
Yield_ton_per_hec	0.01	-0.03	0.09	0.08	0.09	0.01	0.03	0.00	-0.02	0.03	1.00	-0.05	-0.02	-0.01	0.08
Crop_Type_kharif	0.03	0.06	-0.10	-0.07	-0.34	-0.01	0.16	0.43	0.08	-0.00	-0.05	1.00	-0.49	-0.22	-0.48
Crop_Type_rabi	0.01	0.11	-0.08	-0.05	-0.17	0.02	-0.59	-0.60	0.14	0.11	-0.02	-0.49	1.00	-0.17	-0.37
Crop_Type_summer	0.04	0.05	-0.01	0.00	-0.10	-0.02	-0.26	0.24	-0.07	-0.05	-0.01	-0.22	-0.17	1.00	-0.17
Crop_Type_whole year	-0.07	-0.20	0.20	0.13	0.61	-0.00	0.58	-0.01	-0.19	-0.09	0.08	-0.48	-0.37	-0.17	1.00

A photograph of two data analysts in a futuristic control room. One analyst in the foreground wears a white lab coat and blue-tinted glasses, pointing at a large screen displaying various data visualizations like line graphs and heatmaps. Another analyst in the background is also focused on the screens. The room has a modern, industrial feel with large windows showing a city skyline at dusk.

# Exploratory Data Analysis (EDA)



## Summary Statistics

Computed basic descriptive statistics (mean, median, mode, standard deviation) to get an overview of the dataset.

## Univariate Analysis

Analyzed the distribution of individual variables using histograms and box plots to identify patterns and anomalies.

## Bivariate Analysis

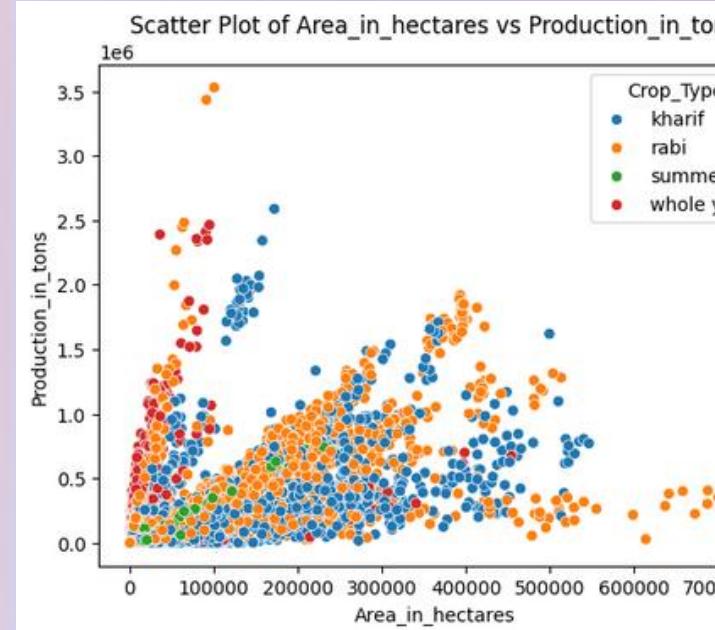
Examined the relationship between two variables using scatter plots and correlation coefficients.

## Multivariate Analysis

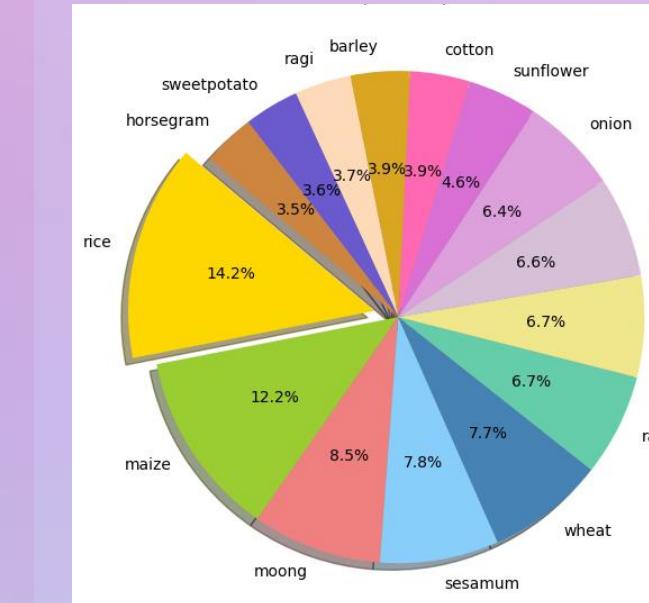
Explored interactions between multiple variables using pair plots and advanced visualization techniques to uncover complex relationships.

# What's EDA ?

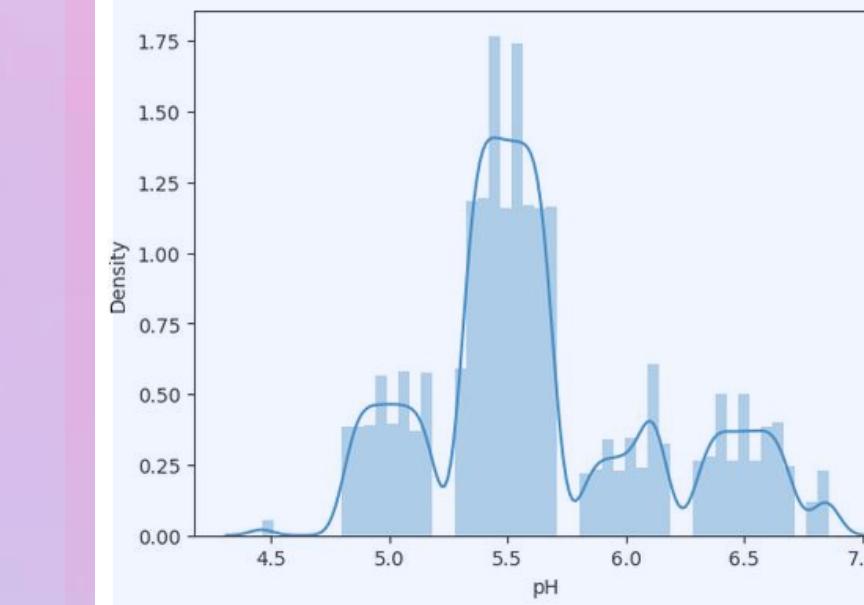
Scatter Plot



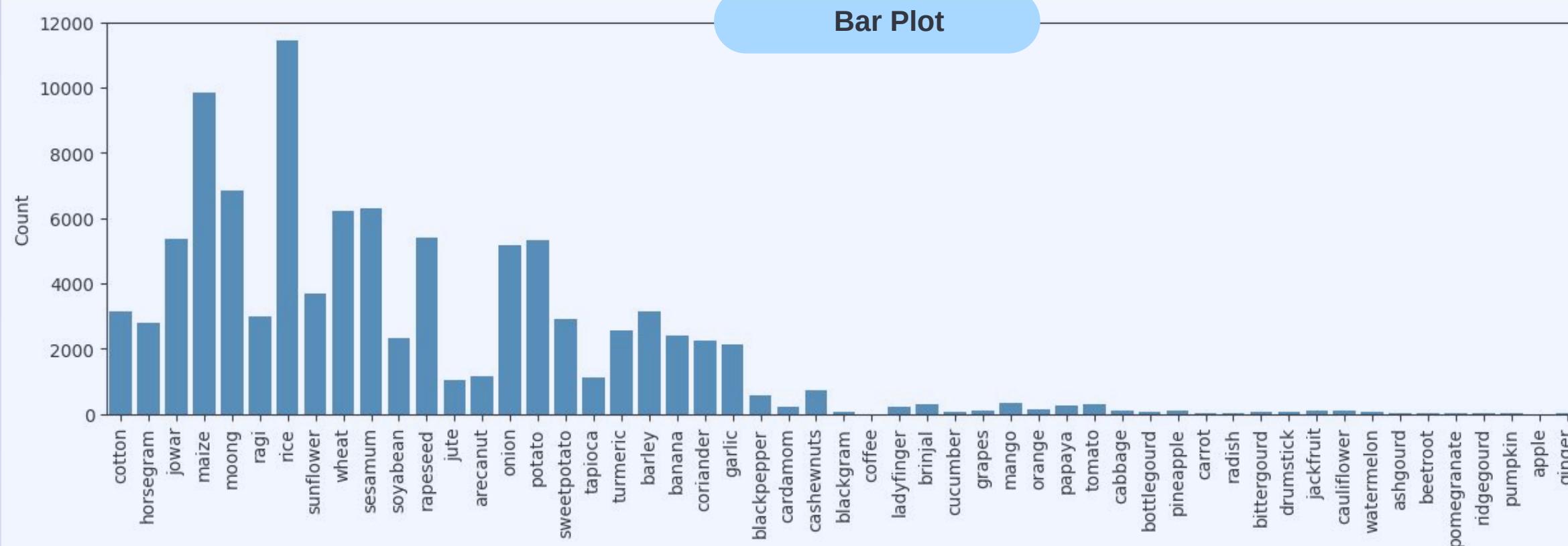
Pie Graph



Bar Graph

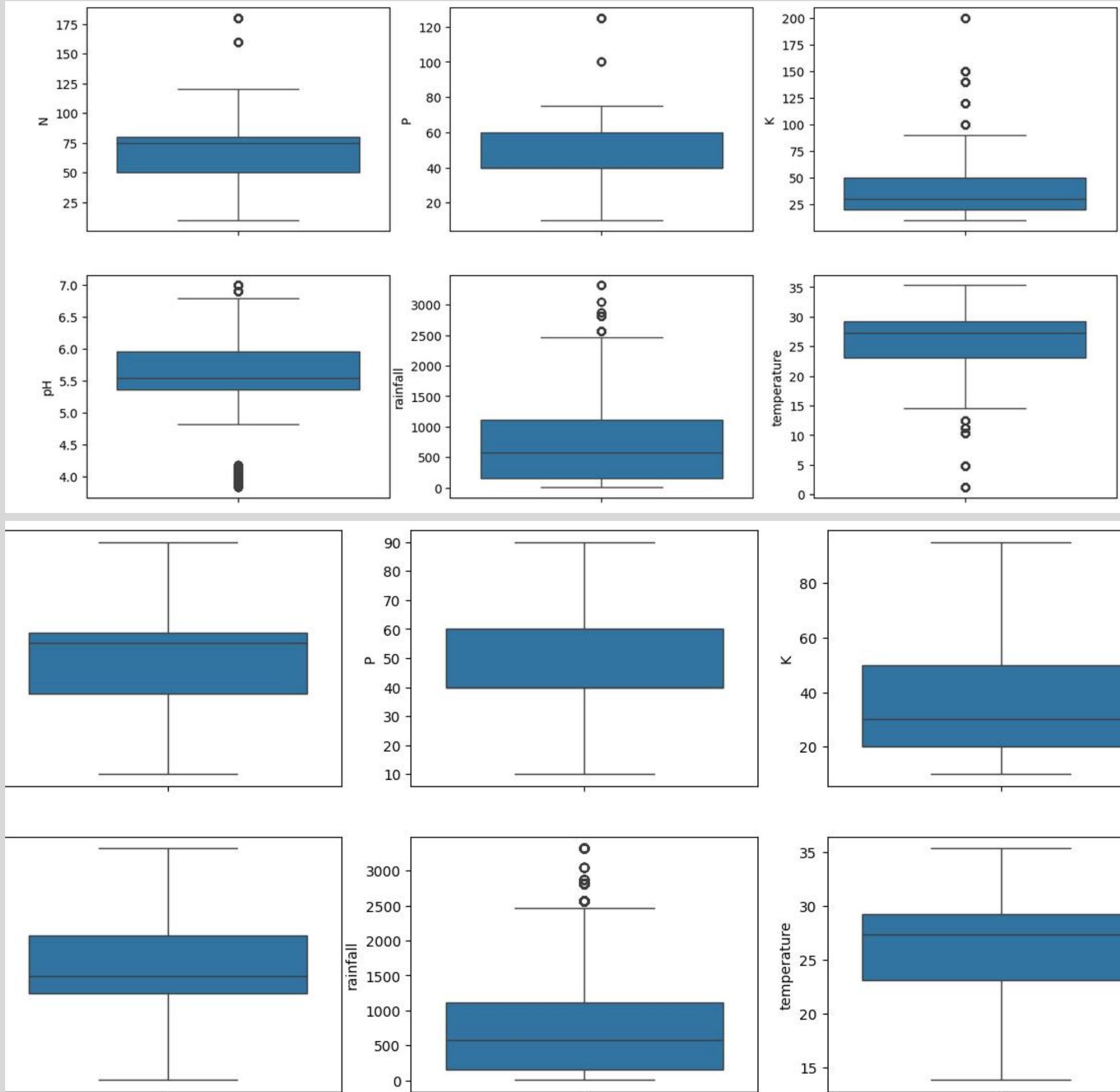


Bar Plot



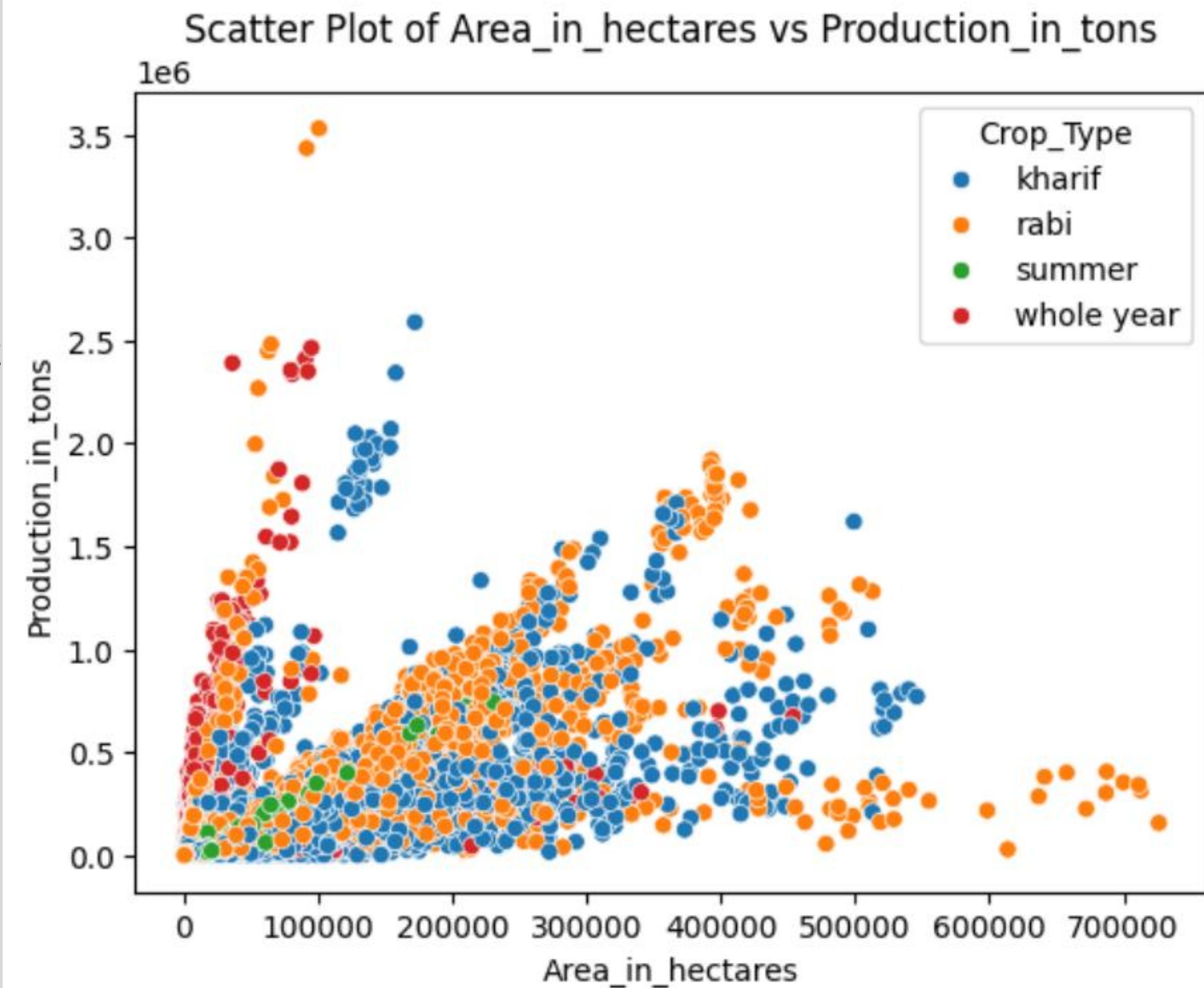
# BOXPLOT

- **Nutrient Levels (N, P, K):** Stable ranges, essential for balanced crop growth.
- **pH:** Slightly acidic to neutral (5.5 - 6.5), suitable for most crops.
- **Rainfall:** Sufficient range (400 - 1500 mm) for crop hydration.
- **Temperature:** Ideal range (15°C - 35°C) for diverse crops.
- **Land Area & Production:** Typical values show moderate land use and production output.
- **Yield:** Most yields are realistic, staying below 1000 tons/hectare.



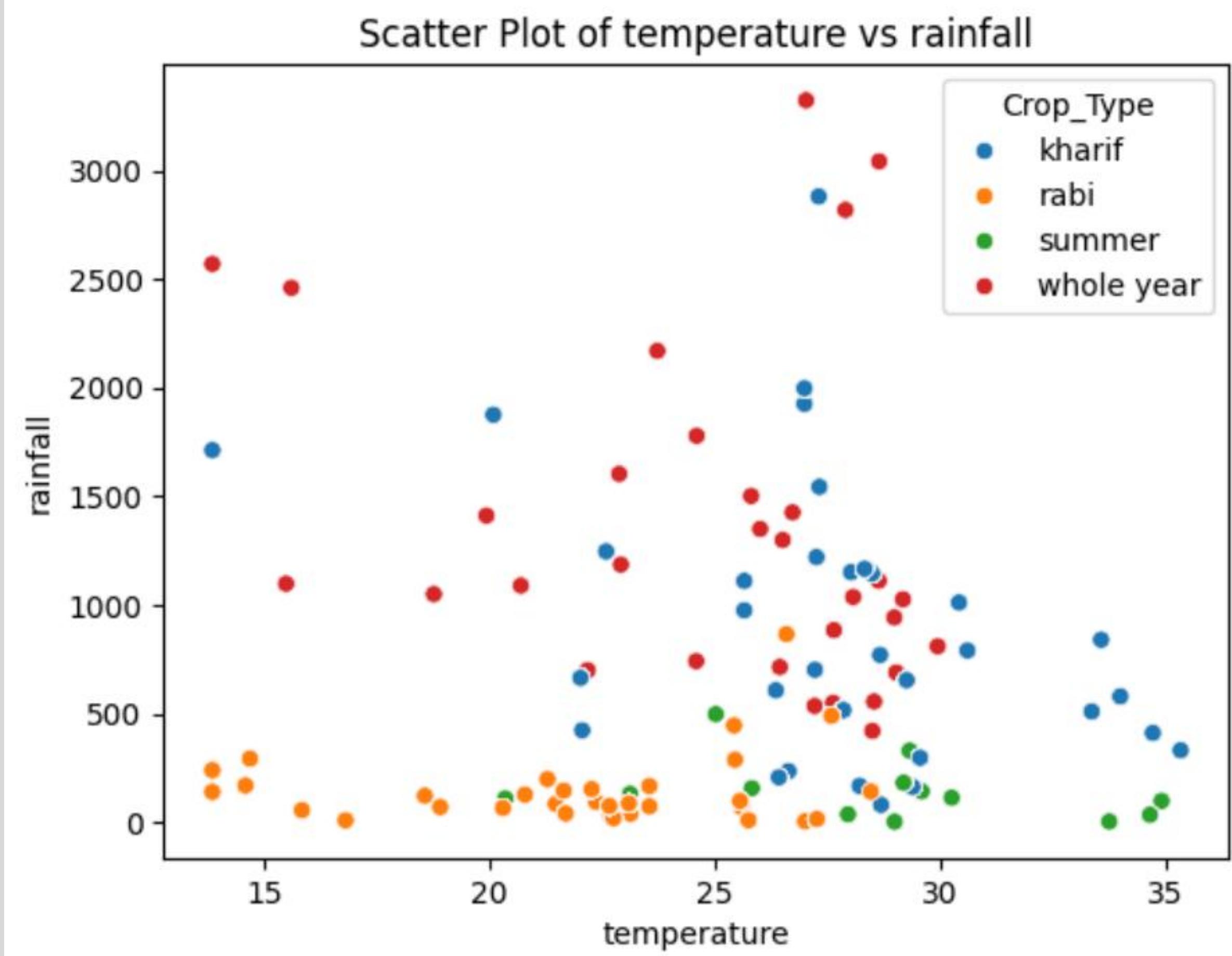
# SCATTERPLOT

- **Kharif Crops:** These crops show high production even with medium-to-large cultivated areas, suggesting a favorable environment or efficient yield for Kharif crops.
- **Rabi Crops:** Rabi crops also show a strong presence, with some points indicating high production relative to the cultivated area.
- **Summer and Whole Year Crops:** These crops have scattered production values and are relatively lower in quantity compared to Kharif and Rabi crops.



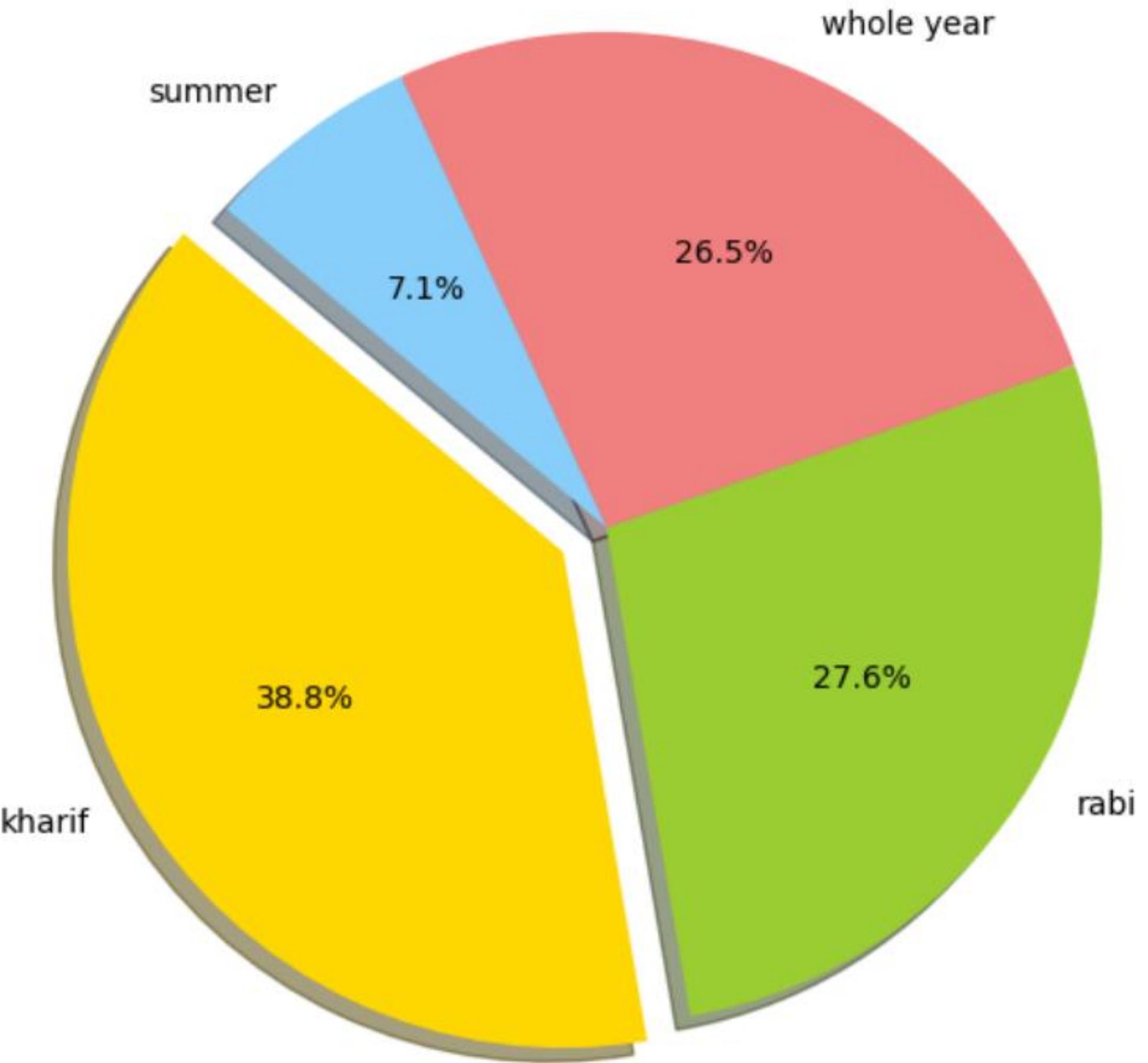
# SCATTERPLOT

- Kharif crops (blue) are generally cultivated with moderate rainfall (500 - 2000 mm) and temperatures around 25-30°C.
- Rabi crops (orange) are grown in lower rainfall (mostly under 500 mm) and cooler temperatures (15-25°C).
- Summer crops (green) thrive in warmer temperatures (25-35°C) but with minimal rainfall.
- Whole-year crops (red) adapt to varying rainfall (0 - 3000 mm) and temperature ranges, showing versatility.



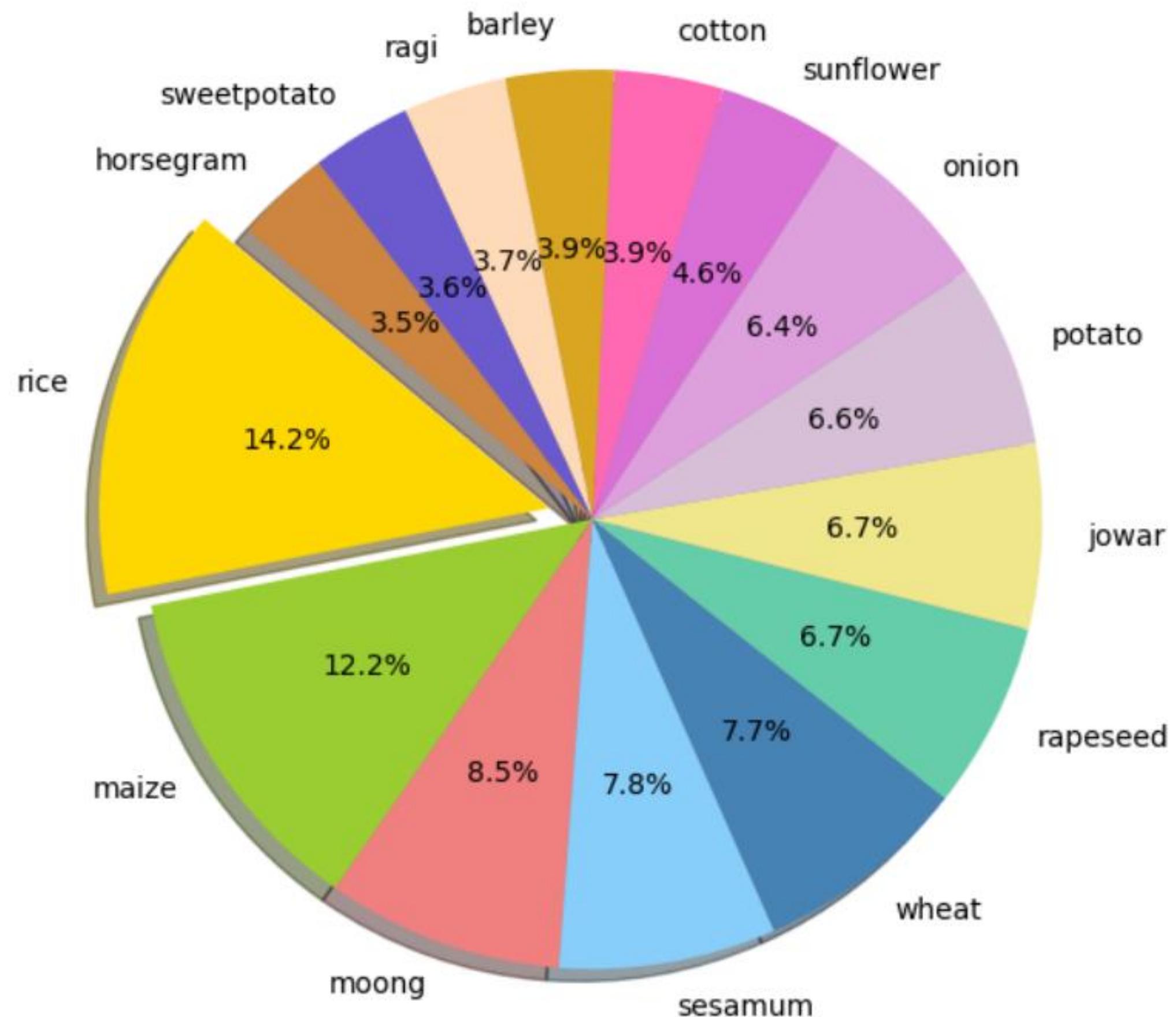
# PIE CHART

- **Leading States:** Uttar Pradesh, Madhya Pradesh, and Karnataka are shown as top contributors to agriculture, with a large percentage of crops cultivated in these regions.
- **Regional Specialization:** States like Telangana, Kerala, and Gujarat show lower cultivation percentages, possibly indicating a focus on specific crops or limited arable land.
- **Implications:** Regional analysis can help identify agricultural strengths and weaknesses of each state, allowing for tailored interventions and resource allocation.



# PIE PLOT

- **Dominant Crops:** The chart shows that rice, maize, and wheat are among the most widely cultivated crops, covering a significant percentage of the agricultural landscape.
- **Less Common Crops:** Certain crops like ragi, horsegram, and sweet potato are grown less frequently, possibly indicating niche cultivation or specific regional demands.
- **Implications:** The variation in crop popularity across regions can provide insights into climate suitability, soil requirements, and market demand for different crops.



# The Importance of Data in Modern Agriculture

## 1 Increased Yields

Optimizing resource allocation and crop management leads to higher crop yields.

## 3 Environmental Sustainability

Efficient resource management and reduced chemical use minimize environmental impact.

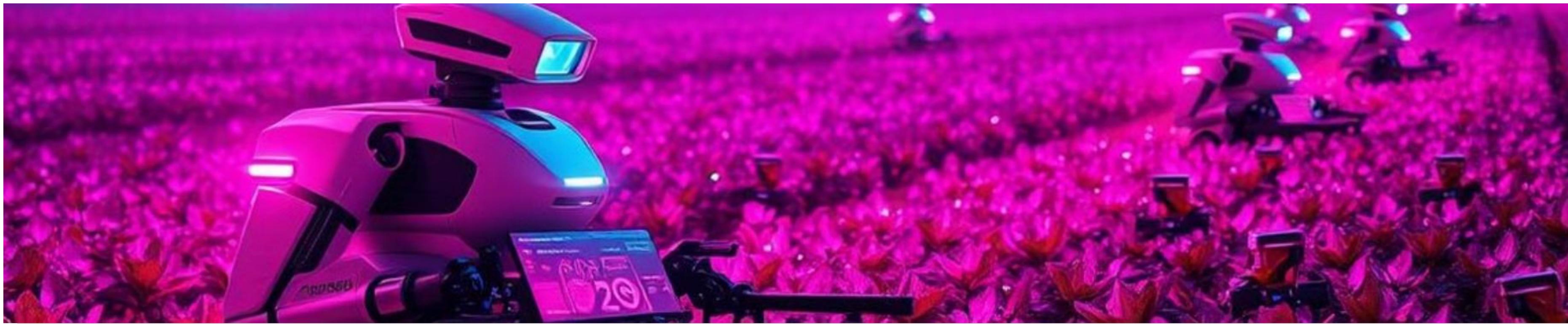
## 2 Improved Profitability

Reduced input costs and increased productivity contribute to higher profits.

## 4 Enhanced Efficiency

Automation and data-driven decisions streamline farm operations and improve efficiency.





# Future Trends in Smart Agriculture

## 1 AI-Powered Robots

Robots equipped with AI will automate tasks such as planting, harvesting, and weeding.

## 2 Precision Livestock Management

Data analytics will play a crucial role in optimizing animal health and production.

## 3 Vertical Farming

Controlled environments for crop production will increase efficiency and reduce land requirements.

# PROJECT GROUP

Sr. No.	Roll No.	Student Name
1	2310303251	ROHIT AGRAWAL
2	2310303159	KRISHNA KUMAR
3	2310303157	KHUSHNAM KHATOON
4	2310303114	DIVYANSHU GANGWAR
5	2310303275	SHAKIR HUSSAIN
6	2310303316	SUPRABH SHARMA
7	2310303307	SNEHA GADE
8	2310303340	VANSH GUPTA
9	2310303296	SHUBHAM RAGHAV
10	2310303170	MANOJ
11	2310302004	MANAS VATS
12	2310303048	ANSHIKA SRIVASTAVA
13	2310303194	MUSKAN PATEL

DCS MENTOR:  
**MANAB DAS**

# THANK YOU !

